Structure-Enhanced Adapter for Self-Supervised Heterogeneous Graph Learning

Fengyu Yan[1,2], **Di Jin**[1*], **Xiaobao Wang**[1,2*], **Qianhua Tang**[1], **Dongxiao He**[1]

[1]Tianjin Key Laboratory of Cognitive Computing and Application, College of Intelligence and Computing, Tianjin University, Tianjin, China

[2]Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ), Shenzhen, China

{fengyuyan, jindi, wangxiaobao, tangqianhua, hedongxiao}@tju.edu.cn

Abstract

Real-world heterogeneous data is commonly modeled as heterogeneous information networks (HINs). Building upon advancements in graph neural networks (GNNs), existing research has significantly progressed in semi-supervised and self-supervised paradigms for heterogeneous GNNs (HGNNs). However, these methods overlook inherent structural deficiencies in raw heterogeneous graphs. We identifies unique structural noise in HINs: missing potential critical edges and multi-relational semantically redundant edges, which force existing HGNNs to learn suboptimal representations on fixed topologies. Crucially, prior limited studies address only partial noise while remaining architecturally entrenched and tightly coupled with specific models. To break this bottleneck, we propose a plug-and-play Heterogeneous graph Structure ADaPter (HSADP) that simultaneously resolves task/model decoupling challenges while accounting for HIN-specific structural properties with with two core components: a dynamic homogeneous subgraph enhancer recovering latent topology across semantic views and a learnable heterogeneous edge discriminator dynamically suppressing redundant edges while collaboratively optimizing semantic graphs. Extensive experiments across multi-domain datasets demonstrate our method's effectiveness and compatibility. The adapter significantly boosts node classification accuracy for multiple SOTA approaches and surpasses specially designed heterogeneous graph structure learning models.

Introduction

In real-world scenarios, data is predominantly heterogeneous, as seen in social networks (Jia et al. 2025), knowledge graphs (Teru, Denis, and Hamilton 2020), citation networks (Hu et al. 2020), and recommendation systems (He et al. 2020) where complex structures and diverse semantic features naturally lend themselves to representation as heterogeneous graphs. Such graphs provide more accurate modeling of these authentic environments. Building upon graph neural network research, state-of-the-art heterogeneous graph neural networks have evolved through two primary paradigms: semi-supervised approaches (Wang et al. 2019; Fu et al. 2020; Hu et al. 2020), and self-supervised/unsupervised methods (Wang et al. 2021b; Mo et al. 2024; Tian et al. 2023; Yan et al. 2025).

*Corresponding Authors.

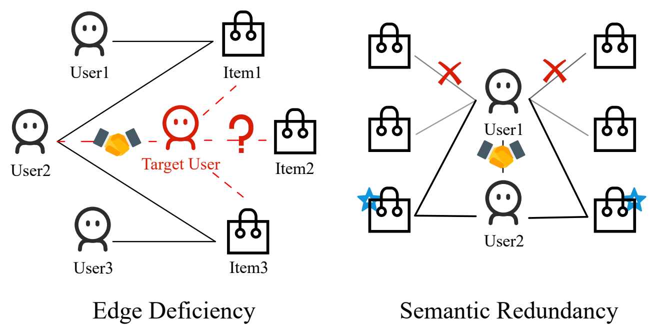

Figure 1: Structural Limitations in HIN: Edge Deficiency: missing potential edges that could optimize message passing (e.g., incomplete user-item interactions in recommendation datasets); Semantic Redundancy: noisy connections dilute effective information flow (e.g., purchases of popular items diminish user differentiation in buying interactions).

However, the existing two paradigms prioritize model design while neglecting raw data optimization, operating under the assumption of fixed heterogeneous graph structures. However, real-world graphs inevitably contain structural noise, such as: (1) human-annotated erroneous edges (Chen, Wu, and Zaki 2020), (2) task-irrelevant redundant (Yu et al. 2020) or incomplete connections (Jin et al. 2020) or (3) adversarially injected anomalous edges from designed attack models (Feng et al. 2024; Wang et al. 2023a, 2025). These imperfections result in suboptimal graph topologies. In homogeneous graph scenarios, extensive research addresses these structural limitations through single-relation adjacency matrix learning, core challenges involve optimizing message propagation across distinct node classes. LDS (Franceschi et al. 2019) employs probabilistic models to generate refined adjacency matrices; GEN (Wang et al. 2021a) integrates KNN (Cover and Hart 1967) clustering to enhance class-aware message passing; ProGNN (Jin et al. 2020) parameterizes adjacency matrices as learnable variables for dynamic optimization. Self-supervised approaches like SUBLIME (Liu et al. 2022b) and GSR (Zhao et al. 2023) introduce topology augmentation perspectives into contrastive learning frameworks. However, these structure learning methods focus exclusively on single-relational adjacency matrices. They are fundamentally limited in modeling multi-relational graphs and fail to capture dependencies between relation types.

Unlike homogeneous graphs, heterogeneous graphs exhibit inherent structural complexity characterized by semantically diverse edges and heterogeneous node types. Consequently, structure learning in such graphs require specialized consideration under multi-relation conditions. This necessity is empirically demonstrated in the two key scenarios presented in Figure 1. (1) **Edge Deficiency:** Mitigation addresses latent multi-relational connections critical for message passing optimization. For instance, in recommendation networks with cold-start users (sparse interactions), recovering potential edges enriches topological completeness and enhances downstream task performance; (2) **Semantic Redundancy:** Filtering high-value signals across heterogeneous edge semantics. For instance, frequent purchases of popular items during recommendation graph construction dilute structural signal integrity, induce systemic noise bias, and degrade user differentiation metrics.

Research on heterogeneous graph structure learning remains limited, existing heterogeneous graph structure learning methods such as HGSL (Zhao et al. 2021), which primarily adds new edges through semantic consistency, they rely on semi-supervised loss co-training and exhibit tight coupling with downstream tasks. Similarly, MV-HGSL (Bing et al. 2023) incorporates structural learning as a contrastive view but remains deeply nested within model architectures, hindering decoupling. This tight coupling necessitates custom structural learning modules for each model, while all aforementioned methods persistently overlook semantic redundancy within heterogeneous information networks.

In summary, *our objective is to design a universal structural learning module that addresses decoupling from downstream tasks while resolving potential specialized structural issues in heterogeneous graphs.* To this end, we developed a heterogeneous graph self-supervised dual-path structural learning adapter. Within this adapter, for Edge Deficiency: We introduce dynamic homogeneous subgraph optimization. Recognizing homophily's critical role in relational subgraphs, our method explicitly models and enhances neighbor homophily connected via meta-paths to discover latent topologies. This ensures learned homogeneous subgraphs not only preserve semantics but also reinforce clustering homogeneity among similar nodes, compensating for missing edges. Consequently, it improves semantic consistency and purity while mitigating noise and ambiguity during subgraph instantiation. For Semantic Redundancy: We design a learnable heterogeneous edge discriminator. This module functions as a multi-relational edge evaluator that assesses the plausibility/importance of each relation-specific edge end-to-end. By explicitly judging the existence of a relation-specific edge between two nodes, it dynamically suppresses noisy edges, enabling fine-grained, relation-aware refinement of the original graph structure. In summary, our key contributions are:

- To the best of our knowledge, we propose the first flexible plug-and-play adapter for self-supervised structure learning in heterogeneous graphs.
- We design a dual-path optimization model specifically addressing edge deficiency and redundancy in multi-relational graphs.
- Extensive experiments demonstrate that our adapter consistently enhances state-of-the-art self-supervised heterogeneous graph methods, achieving competitive performance even on traditional semi-supervised benchmarks.

Related Works

Self-Supervised Heterogeneous Graph Learning

Current self-Supervised approaches are primarily categorized into two types: contrastive-based methods and generative methods. However, these methods typically only consider single-relational scenarios. To address multi-relational scenarios, self-supervised learning methods for heterogeneous graphs have been introduced. DMGI (Park et al. 2020) and HDMI (Jing, Park, and Tong 2021) pre-train models by combining representations across different relations and maximizing mutual information between local and global views. HeCo (Wang et al. 2021b) aligns node representations from different views, selecting positive and negative samples using multi-relational random walks. HGCML (Wang et al. 2023b) leverages the naturally occurring multiple relations in heterogeneous graphs to construct multiple views for contrastive self-supervised training. HGMAE (Tian et al. 2023) employs a generative approach, aligning attribute, structural, and semantic encodings under different relations before and after reconstruction for pre-training. HERO (Mo et al. 2024) pre-trains by aligning node representations from both homogeneous and heterogeneous perspectives within the graph. Nevertheless, all these heterogeneous graph self-supervised learning methods conduct message passing within fixed graph structures, failing to account for structural noise.

Graph Structure Learning (GSL)

Traditional GNN message passing relies exclusively on pre-existing, fixed graph structures. In real-world scenarios, however, these initial fixed structures often contain missing or redundant connections. A series of methods have been proposed to optimize graph structures. As referenced in the OpenGSL benchmark (Zhou et al. 2023), LDS (Franceschi et al. 2019), WSGNN (Lao et al. 2022), and NodeFormer (Wu et al. 2022) learn structures through a joint learning paradigm. GEN (Wang et al. 2021a) and CoGSL (Liu et al. 2022a) iteratively update the structure by alternately training a structure module and a graph encoding module. SUBLIME (Liu et al. 2022b), STABLE (Li et al. 2022) and GSR (Zhao et al. 2023) perform pre-training by constructing structure-augmented views. Nevertheless, all the aforementioned methods target homogeneous graphs and do not consider multi-relational scenarios. Approaches for structure learning in heterogeneous graphs include: GTN (Yun et al. 2019) optimizes the structure by generating variable metapaths via multiple channels and then aggregating them. HGSL

(Zhao et al. 2021) obtains an optimized structure by integrating semantic graph with the feature graph. However, these methods exhibit strong coupling with downstream tasks and models, preventing standalone usage. Moreover, research on structural optimization models for self-supervised scenarios in heterogeneous graphs remains notably lacking.

Problem Statement

Given a heterogeneous graph defined as $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{T}, \mathcal{R}, \mathcal{X}, \tau, \psi)$, where $\mathcal{V} = \{v_1, v_2, ..., v_{|\mathcal{V}|}\}$ is the set of nodes, $\mathcal{E}$ is the set of edges, $\mathcal{T}$ is a set of node types and $\mathcal{R}$ is a set of relation types. $\tau : \mathcal{V} \rightarrow \mathcal{T}$ maps nodes to their types and $\psi : \mathcal{E} \rightarrow \mathcal{R}$ maps edges to their relation types. $\mathcal{X} = \{\mathcal{X}_1, \mathcal{X}_2, ...\mathcal{X}_{|T|}\}$, where $\mathcal{X}_i$ represents the feature for set of node type $i \in \mathcal{T}$.

The self-supervised heterogeneous graph structure learning can be defined as given an optional heterogeneous graph $\mathcal{G}$, the goal of HGSL is to learn the optimal graph structure $\mathcal{G}^* = (\mathcal{V}, \mathcal{E}^*, \mathcal{T}, \mathcal{R}, \mathcal{X}, \tau, \psi)$ to obtain the representation for each predicted node $\mathcal{Z}^* \in \mathbb{R}^{|\mathcal{V}|_{\mathcal{T}_i} \times d}$ for downstream tasks, where $\mathcal{E}^*$ is the refined structure. In a word, the objective of HGSL can be summarized as the following formula:

$$\min_{\theta_f, \theta_{HGSL}} \mathcal{L}_{total} = \mathcal{L}_{task}(\mathcal{Z}, \mathcal{Z}^*) + \mathcal{L}_{HGSL}(\mathcal{E}, \mathcal{E}^*),$$

where θ_f, θ_{HGSL} represents the model for heterogeneous graph encoder and structure model respectfully, $\mathcal{L}_{task}$ stands for the loss for self-supervised task and the $\mathcal{L}_{HGSL}$ is designed for train the structure model.

The Proposed Method

In this section, we present the architecture and design of our structure learning adapter, along with its integration mechanisms with diverse self-supervised graph learning methods.

Overview Framework

The entire architecture comprises two modules: our structure learning module and a pre-training module. Our method decouples from existing pre-training approaches. The optimized graph structure generated by structure learning module is fed into the pre-training module while collaboratively optimizing both components. Our work focuses on the structure learning module. As illustrated in Figure 3, we employ a dual-path design: The homogeneous subgraph enhancer Figure 3.a first extracts multiple semantic-distinct single-relation subgraphs, then dynamically discovers latent relational connections across semantic contexts through feature mapping to enhance semantic consistency and purity. Simultaneously, the heterogeneous edge discriminator Figure 3.b processes node pairs connected by original heterogeneous edges, outputting a pruned structure through edge-wise plausibility assessment. The refined structures from both paths undergo joint optimization with pre-training objectives before transfer to downstream node classification tasks.

Feature Transformation

Based on the characteristics of heterogeneous graphs (with $|\mathcal{T}| + |\mathcal{R}| > 2$), the initial features of nodes often reside in different feature spaces and have inconsistent dimensions. We designed a type-based approach where a dedicated multi-layer perception (MLP) is used for each node category to project the features into a unified dimension. The specific steps are as follows:

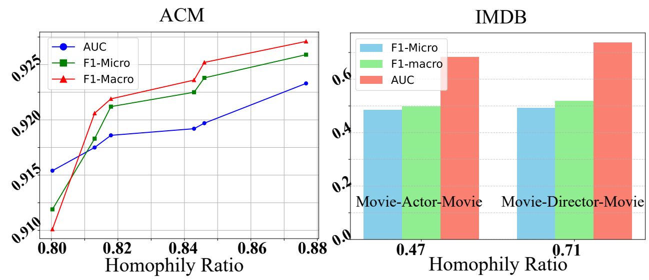

Figure 2: Node classification performance with varying homophily ratios (HR) in semantic homogeneous graphs: These homogeneous graphs are generated from meta-paths in two benchmark datasets. Results demonstrate that higher HR boost classification accuracy.

$$h_i = \sigma(w_{\mathcal{T}_i} \times x_i + b_{\mathcal{T}_i}), \tag{1}$$

where $w_{\mathcal{T}_i}$ and $b_{\mathcal{T}_i}$ is the weight and bias parameters for node type $\mathcal{T}_i$, σ is the activation function and $h_i \in \mathbb{R}^d$ is the representation for node i.

Homogeneous Subgraph Enhancer

This section introduces our proposed method to address the problem of discovering potential beneficial edges and identifying structural augmentations with positive gains to mitigate information incompleteness.

Meta-path-based random walks are a common technique for handling diverse semantics in heterogeneous graphs (Wang et al. 2019). We first extract homogeneous graphs $\{\mathcal{G}_{P_1}, \mathcal{G}_{P_2}..., \mathcal{G}_{P_i}\}$ under different semantics from the original heterogeneous graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{T}, \mathcal{R}, \mathcal{X}, \tau, \psi)$ by leveraging pre-defined meta-paths $\{\mathcal{P}_1, \mathcal{P}_2..., \mathcal{P}_i\}$, each meta-path $\mathcal{P}_i$ can be defined as:

$$\mathcal{P}_i = t_1 \xrightarrow{r_1} t_2 \xrightarrow{r_2} t_3 \cdots t_{k-1} \xrightarrow{r_k} t_k,$$

where $t_i \in \mathcal{T}$ denotes node types and $r_i \in \mathcal{R}$ denotes relation types. The adjacency matrix $\mathbf{A}_{\mathcal{P}_i}$ for semantics homogeneous graph is computed as:

$$\mathbf{A}_{\mathcal{P}_i} = \mathbf{A}_{R_1} \mathbf{A}_{R_2} \cdots \mathbf{A}_{R_k}, \tag{2}$$

where $\mathbf{A}_{R_i} \in \mathbb{R}^{|\mathcal{V}_{\mathcal{T}_i}| \times |\mathcal{V}_{\mathcal{T}_{i+1}}|}$ is the relation-specific adjacency matrix derived from the original topological semantics. We observe that these derived single-relational graphs exhibit varying homophily ratios $HR_{G_{p_i}}$ (the proportion of edges connecting nodes of the same class within a subgraph) which can be defined as:

$$HR_{\mathcal{P}} = \frac{1}{|\mathcal{E}_{\mathcal{P}}|} \sum_{(u,v) \in \mathcal{E}_{\mathcal{P}}} \mathbb{I}(y_u = y_v), \tag{3}$$

where y_i denotes node class, $|\mathcal{E}_{\mathcal{P}}|$ denotes edge count in meta-path subgraph and indicator function $\mathbb{I}$ (1 when

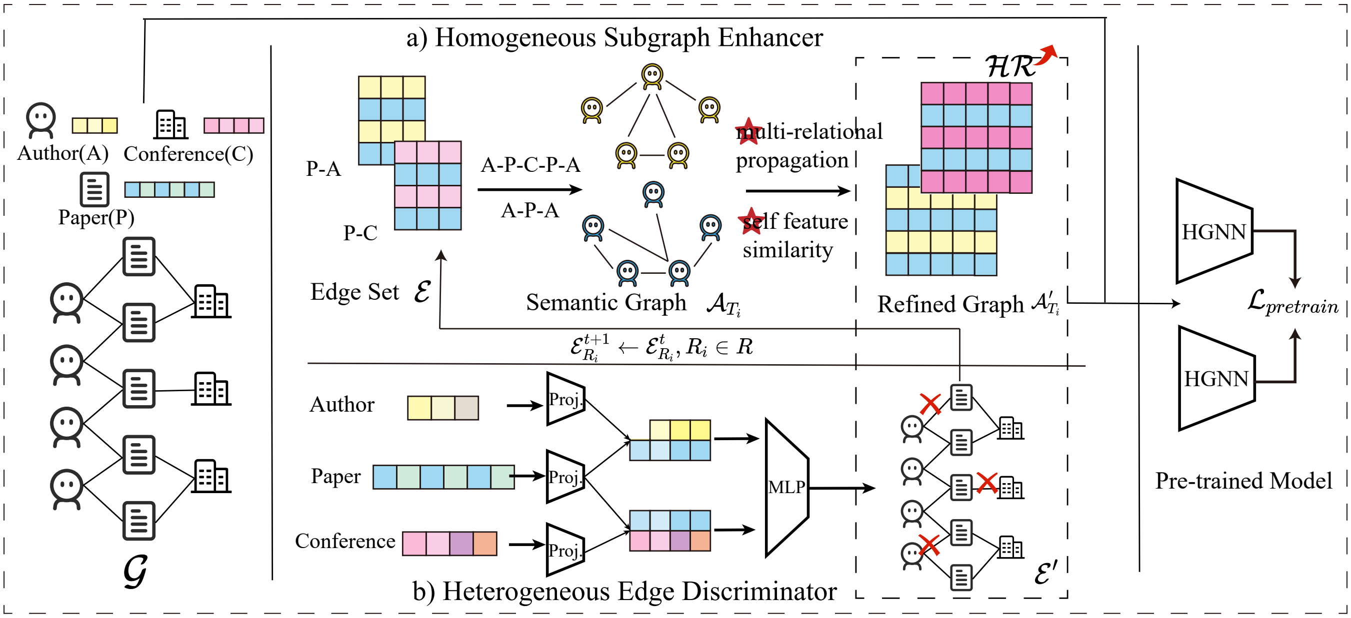

Figure 3: Framework of the Proposed Structure Adapter: Our method optimizes original graph structures for heterogeneous graph pre-training through a dual-path architecture that concurrently addresses potential structural deficiencies (Homogeneous Subgraph Enhancer) through two complementary mechanisms and semantic redundancy (Heterogeneous Edge Discriminator), collectively rectifying structural oversights in graph pre-training and enabling end-to-end joint optimization.

$y_u = y_v$, else 0). Furthermore, classification performance improves when using semantic homogeneous graphs with higher HRs as depicted in Figure 2. The semantic homogeneous graphs are generated from the ACM dataset using six distinct meta-paths and the IMDB dataset through two meta-paths. These results confirm that homophily enhancement leads to measurable performance gains in target tasks. Accordingly, we design two specialized homophily increasing mechanisms: feature similarity propagated neighborhood expansion and multi relation constrained neighborhood expansion. Given the obtained set of semantic homogeneous graphs $\{\mathcal{G}_{p_1}, \mathcal{G}_{p_2} ..., \mathcal{G}_{p_i}\}$, we first compute cross-type node similarity, expressed as:

$$\text{sim}(u,v) = \begin{cases} \frac{\mathbf{h}_u \cdot \mathbf{h}_v}{\|\mathbf{h}_u\| \|\mathbf{h}_v\|} & \text{if } \tau(u) = \tau(v) \\ 0 & \text{otherwise} \end{cases}, \tag{4}$$

where h_i is the feature vector of node i, and $\tau(i)$ denotes node type mapping function. At this stage, we obtain intra-type node similarity matrices S_{T_i} for each type T_i, categorized into target node similarity matrix S_{T_t}, where T_t is the type of target node and non-target node similarity matrices $\{S_{T_i}, i \in |T| \;\; \& \;\; i \neq t\}$. For self feature similarity neighborhood expansion: within each node class, we set a threshold η to connect node pairs with high similarity in the predictive node similarity matrix, which can be computed as:

$$\tilde{\mathcal{A}}_{\mathrm{T}_t} = \left\{ (v_i, v_j) \mid v_i, v_j \in \mathcal{V}^{(T_t)}, sim(v_i, v_j) \geq \eta \right\}, \tag{5}$$

where $\mathcal{B}_{\mathrm{T}_t}$ denotes the potential structure, discovers beneficial edges and increases the homophily ratio of the single-relational graph from self neighbor expansion.

For dynamic neighbor expansion based on multi-relational propagation: By capturing correlations between adjacent nodes, we identify pairs of non-target nodes b_k, b_l with similarity exceeding threshold δ, and target nodes a_i, a_j that connect to these non-target nodes with sufficient frequency. A new edge is formed between a_i and a_j when:

$$\tilde{\mathcal{A}}_{\mathrm{T}_{i \neq t}} = \left\{ (a_i, a_j) \middle| \begin{array}{l} \exists b_k, b_l \in \mathcal{V}^{(T_i)} \\ sim(b_k, b_l) \geq \delta \\ \deg(a_i \to b_k) \geq \gamma \\ \deg(a_j \to b_l) \geq \gamma \end{array} \right\}, \tag{6}$$

where $\deg(a_i \to b_k) = \sum_{e \in \mathcal{E}} \mathbb{I}[(a_i, b_k) \in \mathcal{E}]$ and γ is the count of existing edges. This establishes edges between target nodes that related to similar non-target nodes, enhancing structural completeness. Finally, we integrate the potential structures from both methods with the original structure using attention mechanism, and apply sparsification operations to obtain the updated structure under the homogeneous subgraph enhancement path, which can be computed as:

$$\mathcal{A}'_{T_i} = \text{Sparsify}(\beta_1 \odot \mathcal{A}_{T_t} + \beta_2 \odot \sum_{i \in T \setminus T_t} \mathcal{A}_{T_i}). \tag{7}$$

To reduce computational overhead, we employ top-K sparsification, where the sparsification function is defined as:

$$\text{Sparsify}(\mathbf{A})_{ij} = \begin{cases} \mathbf{A}_{ij} & \text{if } \mathbf{A}_{ij} \in \text{topk}(\mathbf{A}_{i:}) \\ 0 & \text{otherwise} \end{cases}, \tag{8}$$

$$\mathcal{A}'_{P_i} = \alpha \odot \mathcal{A}_{P_i} + \beta \odot \mathcal{A}'_{T_i}, \tag{9}$$

where α, β are the learnable parameters, $\mathcal{A}_{\mathcal{P}_\rangle}{}'$ is the refined adjacency matrix for each homogeneous graph generated by meta-path P_i.

Heterogeneous Edge Discriminator

In the preceding section, we identify potential edges that could augment graph connectivity. However, for the second problem, indiscriminate edge addition introduces redundant information that dilutes critical structural signals and introduces undesirable noise or bias into semantic representations. To address this, we propose a multi-relational heterogeneous edge discriminator that evaluates edge necessity across distinct relation types. This module operates through sequential operations. Firstly, for each candidate edge e_{ij} connecting nodes v_i, v_j under relation r, the discriminator computes necessity probability p_{ij} by:

$$p_{ij} = \phi_r \left(\mathbf{h}_i \parallel \mathbf{h}_j \parallel \mathbf{h}_{rel} \right), \tag{10}$$

where ϕ_r denotes a relation-specific MLP, $\mathbf{h}_{rel}$ is the relation embedding and $||$ is concatenation. This yields probabilistic edge scores across all relation types. The relational encoding is optional and can utilize pre-trained methods such as Metapath2Vec (Dong, Chawla, and Swami 2017). We incorporate the edge necessity probabilities into graph construction, where all original edges remain learnable parameters. This updated topology serves dual purposes: providing the foundation for heterogeneous graph pre-training and enabling iterative homogeneous subgraph updating through dynamic adjacency matrices. The update mechanism follows:

$$\mathcal{E}_{R_i}^{t+1} \leftarrow \mathcal{E}_{R_i}^{t}, R_i \in R, \tag{11}$$

where $\mathcal{E}_{R_i}^{t}$ means the relation R_i in t iteration. Thus, we obtain the structurally enhanced dual-path output for subsequent heterogeneous graph pre-training input.

Experiments

To demonstrate the validity of our approach, we conduct comprehensive experiments across five benchmark datasets. The experimental methodology is sectionally organized as follows: (1) configuration details (parameter settings, baseline specifications); (2) quantitative results with comparative analysis demonstrating model efficacy.

Experimental Setup

Datasets To evaluate the effectiveness of our proposed method, we conduct experiments on five widely-used heterogeneous graph datasets: **ACM** (Hu et al. 2020), **DBLP** (Hu et al. 2020), **Freebase** (Lv et al. 2021), **Yelp** (Zhao et al. 2021), and **AMiner** (Hu, Fang, and Shi 2019). These datasets span various domains and exhibit diverse node and relation types, making them representative benchmarks for heterogeneous graph representation learning.

- **ACM** is constructed from a subset of papers published in ACM conferences. It contains four types of nodes: papers, authors, terms, and subjects, and two main types of edges: *paper-to-author* and *paper-to-subject*.
- **DBLP** is a bibliographic dataset containing information about authors, papers, venues, and terms. The graph is built with various node types (e.g., author, paper, venue) and relation types (e.g., author-to-paper, paper-to-venue).
- **Freebase** is a large-scale knowledge graph consisting of various types of entities such as books, films, and organizations. It contains rich semantic relations, e.g., *book-to-film*, *organization-in-film*.
- **Yelp** is built from business review data. It includes users, businesses, tips, and reviews as nodes, with edges like *user-writes-review* and *review-about-business*.
- **AMiner** is an academic network composed of nodes such as authors, papers, and venues. It includes edges like *paper-written-by-author* and *author-writes-paper*.

Dataset	#Nodes	#Types	#Edges	#Types	TargetNode
ACM	8,994	3	25,922	4	paper
DBLP	26,128	4	296,563	3	author
Freebase	180,098	8	1,057,688	36	book
Yelp	3,913	4	72,132	6	business
AMiner	55,783	3	153,676	4	paper

Table 1: The statistics of the heterogeneous graph datasets

Baselines To demonstrate the effectiveness of our approach, we compare against four categories of baselines: (1) Traditional semi-supervised homogeneous graph learning GCN (Kipf and Welling 2017), GAT (Veličković et al. 2018), designed for single relation; (2) Semi-supervised heterogeneous graph learning HAN (Wang et al. 2019), HGT (Hu et al. 2020), designed for multi-relations, structure-only methods Mp2vec (Dong, Chawla, and Swami 2017); (3) Heterogeneous graph structure learning HGSL (Zhao et al. 2021), leveraging label information for structural refinement and representation enhancement; (4) Heterogeneous graph self-supervised learning, focusing on unlabeled scenarios, including: contrastive view-based approaches HeCo (Wang et al. 2021b), HERO (Mo et al. 2024), HGCML (Wang et al. 2023b), reconstruction-based techniques HGMAE (Tian et al. 2023). Our adapter is embedded into the aforementioned heterogeneous graph pre-training models for joint training.

Parameter Settings Comprehensive experimental configurations and evaluation protocols are delineated herein. To rigorously validate pre-training efficacy, we implement a methodology where 10-shot and 20-shot instances per class are randomly sampled across datasets to construct distinct training sets. Each experimental scenario undergoes five independent training-testing cycles to ensure equitable comparison, with final results aggregated as the mean value of trials and variance reported to quantify stability. Evaluation employs three established metrics: Macro-F1 (class-balanced performance), Micro-F1 (instance-weighted performance), and AUC (ranking capability). For all baseline models, we strictly preserve their original published configurations and parameter settings without modification. Our structural adapter implementation involves several critical hyper-parameters: meta-path selection (utilizing dataset-specific prior knowledge for optimal path identification), semantic graph similarity thresholds $\eta = 0.6$, cross-type node

Method	ACM		Freebase		AMiner		DBLP		Yelp	
	Ma-F1	Mi-F1	Ma-F1	Mi-F1	Ma-F1	Mi-F1	Ma-F1	Mi-F1	Ma-F1	Mi-F1
10 shot										
GCN	72.90±0.2	73.16±0.3	31.01±0.2	34.42±0.3	54.30±1.1	58.27±2.3	69.05±1.8	71.09±2.7	69.63±3.2	67.12±2.9
GAT	70.72±0.3	71.06±0.5	30.93±1.2	34.26±1.8	55.97±2.5	59.26±3.2	75.39±3.2	76.06±1.6	69.52±2.4	64.81±1.6
HAN	82.67±0.2	82.87±0.4	33.14±1.6	35.16±1.7	60.91±0.8	65.84±1.0	84.93±1.3	84.16±2.9	80.38±1.2	83.01±0.6
HGT	76.70±0.1	76.49±0.2	34.57±2.6	38.07±2.3	64.33±0.7	68.81±0.2	77.48±2.1	80.16±2.8	78.99±2.2	74.27±1.1
HGSL	82.89±0.2	85.65±0.3	40.23±0.8	42.96±0.6	65.79±0.7	70.61±0.9	85.44±0.3	86.83±0.3	79.36±0.4	84.94±0.8
HeCo	87.76±0.2	86.83±0.2	36.55±2.2	45.02±2.0	67.31±1.4	76.46±1.8	89.56±0.3	**90.36±0.4**	82.49±0.4	83.24±0.2
+HSADP	**90.39±0.3**	**90.22±0.2**	40.36±4.0	46.66±2.5	68.58±0.2	79.41±1.0	88.41±0.3	89.18±0.4	83.89±0.2	84.74±0.5
HGCML	85.36±0.4	85.44±0.8	42.32±0.7	47.89±0.6	77.35±0.5	88.26±0.2	87.38±0.1	86.76±0.2	84.87±0.7	87.65±0.4
+HSADP	87.64±0.5	86.94±0.3	45.49±0.7	50.76±0.6	78.62±0.7	**88.92±0.6**	89.84±0.3	87.66±0.3	85.90±0.6	88.51±0.8
HGMAE	88.10±0.1	87.38±0.1	47.33±1.3	52.75±1.8	68.32±0.1	77.01±0.1	88.25±0.1	87.27±0.1	80.30±0.7	83.87±0.6
+HSADP	89.42±0.1	88.33±0.1	**52.64±1.0**	**58.37±0.9**	70.69±0.1	78.72±0.1	**90.26±1.1**	89.08±0.9	83.81±0.7	87.49±0.7
HERO	85.28±0.4	85.37±0.3	43.52±0.3	45.29±0.6	79.04±0.2	87.98±0.3	86.28±0.2	87.89±0.1	86.91±0.3	86.54±0.4
+HSADP	87.28±0.3	85.87±0.5	48.23±0.5	46.77±0.2	**79.23±0.1**	88.43±0.4	89.18±0.6	89.76±0.3	**89.48±0.5**	**89.73±0.4**
20 shot										
GCN	76.70±0.1	76.77±0.1	46.89±0.2	46.39±0.1	60.21±0.7	66.47±0.3	82.29±0.1	82.68±0.1	75.96±0.2	79.57±0.2
GAT	73.95±1.2	70.42±3.5	42.62±0.2	44.84±0.1	57.32±1.3	62.93±1.2	83.40±0.1	83.47±0.1	76.05±0.6	79.32±0.0
HAN	85.61±1.2	85.56±1.3	50.26±0.1	49.97±0.1	66.89±0.5	72.16±0.4	86.89±0.6	84.73±1.1	84.22±0.1	83.10±0.1
HGT	87.60±2.8	87.96±5.4	54.87±0.2	51.55±0.4	65.23±0.3	69.48±0.5	86.46±1.1	83.68±0.3	84.13±0.6	86.63±0.1
Mp2vec	68.44±0.1	67.02±0.1	44.21±0.1	50.71±0.1	60.78±0.5	65.32±0.8	79.01±0.0	75.94±0.0	75.19±0.0	78.11±0.0
HGSL	88.42±0.8	87.12±0.5	58.32±0.7	55.83±0.4	68.43±0.4	76.34±0.2	88.49±0.2	87.73±0.3	85.28±0.9	86.73±0.7
HeCo	87.39±0.4	86.71±0.6	57.84±0.4	60.12±0.6	71.39±0.4	78.06±0.6	90.39±0.4	90.06±0.6	85.32±0.4	87.22±0.2
+HSADP	90.10±0.1	89.86±0.1	59.90±1.6	62.94±1.7	73.46±0.9	79.47±1.1	91.38±0.2	91.21±0.2	86.92±0.8	89.44±0.2
HGCML	90.43±0.2	90.26±0.6	61.72±0.1	64.11±0.5	71.82±0.4	80.47±0.7	91.82±0.4	92.24±0.6	90.34±0.4	90.24±0.7
+HSADP	91.28±0.4	90.14±0.3	62.22±0.5	65.87±0.6	73.42±0.2	81.62±0.2	92.46±0.8	93.18±0.4	92.04±0.6	91.58±0.7
HGMAE	90.38±0.4	90.06±0.6	61.42±0.1	64.61±1.5	72.52±0.4	80.07±0.7	91.52±0.6	92.07±0.6	90.30±0.2	90.72±0.6
+HSADP	**91.38±0.5**	**90.94±1.1**	62.62±1.1	65.77±0.8	73.76±0.2	81.69±0.8	92.36±1.1	93.08±0.9	**91.94±0.7**	**91.48±0.6**
HERO	87.68±0.6	87.60±0.2	61.22±0.3	65.23±0.6	76.85±0.1	85.94±0.2	91.35±0.3	92.38±0.2	88.70±0.1	87.98±0.3
+HSADP	91.23±0.8	90.35±0.4	**63.42±0.5**	**66.23±0.2**	**79.46±0.2**	**88.42±0.1**	**93.48±0.2**	**94.38±0.1**	90.58±0.2	90.43±0.3

Table 2: Performance presented with the best bolded, improvements underlined, and '+' denoting adapter integration

structural affinity thresholds $\gamma = 6$ and $K = 10$ in sparsification function. When co-training with meta-models, our adapter naturally integrates contrastive learning approaches. In meta-path views, semantically augmented views are employed, while attention-based methods utilize heterogeneous edge discriminators. For generative approaches, both semantically enhanced graphs and heterogeneous edge discriminators are embedded within structure-aware reconstructions, with message passing operating on the updated architectures. Learning rates are fixed at 8e-4 for pretrained models and 0.01 for the classifier. Optimization utilizes the Adam optimizer (PyTorch implementation). with all trials conducted on NVIDIA RTX 3090 graphics cards.

Performance Evaluation

This section utilizes the established experimental configuration to verify the effectiveness of the heterogeneous structure adapter for node classification, assess the significance of individual modules, and evaluate the adapter's sensitivity to critical hyperparameters.

Node Classification In the node classification task, we employ five heterogeneous information network datasets spanning multiple domains including business and academia, etc. Our model is embedded as a plugin into the heterogeneous graph pre-training models mentioned in the baselines and is jointly trained during the pre-training phase. To investigate the significance of our model within the pre-training framework, we conduct comparative experiments for both 10-shot and 20-shot settings. As evidenced by the experimental results in Table 2, we draw the following conclusions: 1) Our structural adapter demonstrates consistent improvements in node classification performance across nearly all benchmark datasets for state-of-the-art heterogeneous graph pre-training models. These results highlight the value of structural learning in self-supervised heterogeneous graph representation and confirm the adapter's effectiveness; 2) While baseline pre-trained models demonstrate sustained efficacy in some

Dataset	Metric	HAN	+HSADP	HGT	+HSADP
ACM	Ma-F1	85.61±1.2	87.06±0.1	87.60±2.8	**88.75±0.1**
	Mi-F1	85.56±1.3	87.09±0.3	87.96±5.4	**88.95±0.1**
	AUC	90.25±0.2	**96.98±0.1**	89.86±0.1	92.09±0.1
DBLP	Ma-F1	86.89±0.6	**88.34±0.2**	86.46±1.1	86.70±0.3
	Mi-F1	84.73±1.1	**88.70±0.0**	83.68±0.3	84.01±0.5
	AUC	92.89±0.1	**94.86±0.1**	88.65±0.1	89.67±0.2
Yelp	Ma-F1	84.22±0.1	85.04±0.6	84.13±0.6	**86.01±0.1**
	Mi-F1	83.10±0.1	**86.78±0.0**	86.63±0.1	84.01±0.1
	AUC	84.40±0.1	**85.60±0.1**	82.44±0.1	84.62±0.0
Freebase	Ma-F1	50.26±0.1	53.01±0.1	54.67±0.2	**54.78±0.1**
	Mi-F1	49.97±0.1	**55.42±0.1**	51.55±0.4	54.04±6.2
	AUC	69.45±0.1	**74.62±0.1**	68.33±0.0	73.60±0.2

Table 3: Performance comparison across different HGNN methods with best bolded, improved underlined and "+" indicates adapter plugged in.

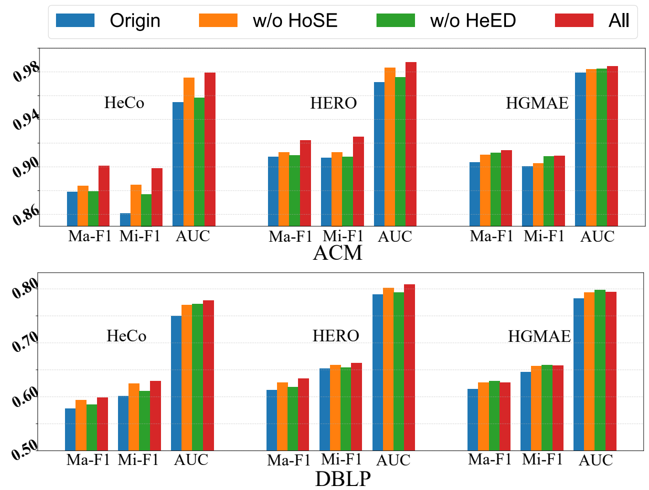

Figure 4: Ablation study

datasets across diverse few-shot data scales, they suffer significant performance deterioration in others. Our adapter universally boosts baseline performance and effectively alleviates sharp declines, with the most notable improvements observed at the 10-shot setting; 3) Our adapter demonstrates consistent performance superiority over supervised heterogeneous structure learning methods, confirming both its compatibility with diverse graph architectures and empirical efficacy; 4) Pre-training approaches demonstrate significant performance advantages over conventional homogeneous and heterogeneous graph learning methods. Notably, as evidenced by the results in Table 3, our method further enhances the performance of standard heterogeneous graph models when integrated with them.

Ablation Study To comprehensively investigate the impact of individual components within our model, we conduct ablation studies on critical modules: (1) w/o HoSE: removing the HOmogeneous Subgraph Enhancement module from the structure adapter; (2) w/o HeED: excluding the HEterogeneous Edge Discriminator. These ablations are evaluated across two datasets using three heterogeneous graph pre-training methods spanning two categories. Results in Figure 4 demonstrate that both modules significantly enhance structural representations and improve model performance. The dual-path approach exhibits complementary effects, enabling joint structural optimization. Notably, for the HERO method, incorporating meta-path information substantially boosts model capability.

Hyper-parameters Analysis We conduct a rigorous ablation study on three core hyper-parameters of our structural adapter: (1) homogeneous similarity threshold η for self-feature based neighbor selection, (2) cross-type correlation threshold γ for multi-relational propagation, and (3) top-K in sparsification process. Empirical results on ACM and DBLP datasets using two pre-trained model illustrated in Figure 5, which indicate that model performance is sensitive to the similarity threshold setting, with both overly lenient and strict thresholds adversely affecting accuracy due to fundamental constraints in node correlation distributions. Our sparsification process integrated with heterogeneous edge discrimination successfully addresses the redundancy issues arising from aggressive node selection.

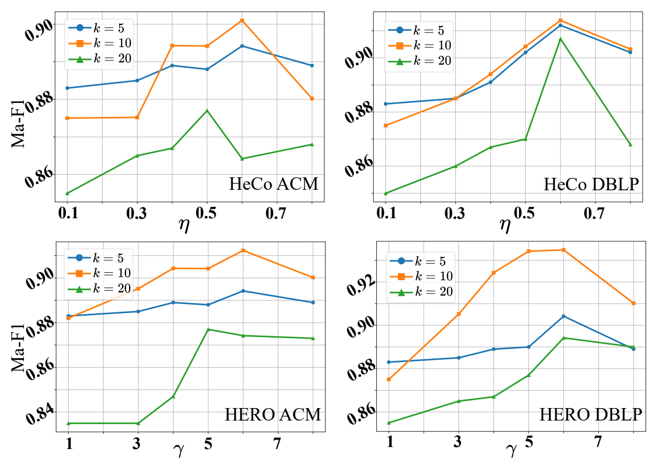

Figure 5: Impact of hyper-parameters

Conclusion

This study focuses on self-supervised structural learning for heterogeneous graphs, addressing the gap in structural modeling within existing heterogeneous graph pretraining methods. Our approach serves as a structural adapter, seamlessly integrating with established pretraining frameworks. Experiments demonstrate its effectiveness and generality. Currently, our method is constrained to structural enhancements within single datasets. Future work will extend this by: advancing multi-dataset joint training mechanisms for structural learning, exploring broader scenarios including zero-shot transfer and leveraging structural understanding from large language models for enhanced adaptation.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2023YFC3304503), the National Natural Science Foundation of China (No. 62422210, 92370111, 62302333, 62272340, 62276187), the Open Research Fund from Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ) (No. GML-KF-24-16) and Hebei Natural Science Foundation (No. F2024202047).

References

Bing, R.; Yuan, G.; Meng, F.; Wang, S.; Qiao, S.; and Wang, Z. 2023. Multi-view Contrastive Enhanced Heterogeneous Graph Structure Learning. *Int. J. Softw. Informatics*, 13(4): 417–447.

Chen, Y.; Wu, L.; and Zaki, M. J. 2020. Iterative Deep Graph Learning for Graph Neural Networks: Better and Robust Node Embeddings. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020*.

Cover, T. M.; and Hart, P. E. 1967. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory*, 13(1): 21–27.

Dong, Y.; Chawla, N. V.; and Swami, A. 2017. metapath2vec: Scalable Representation Learning for Heterogeneous Networks. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 135–144.

Feng, B.; Jin, D.; Wang, X.; Cheng, F.; and Guo, S. 2024. Backdoor attacks on unsupervised graph representation learning. *Neural Networks*, 180: 106668.

Franceschi, L.; Niepert, M.; Pontil, M.; and He, X. 2019. Learning Discrete Structures for Graph Neural Networks. In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019*, 1972–1982.

Fu, X.; Zhang, J.; Meng, Z.; and King, I. 2020. MAGNN: Metapath Aggregated Graph Neural Network for Heterogeneous Graph Embedding. In *WWW '20: The Web Conference 2020, Taipei, Taiwan, April 20-24, 2020*, 2331–2341.

He, X.; Deng, K.; Wang, X.; Li, Y.; Zhang, Y.; and Wang, M. 2020. LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 639–648.

Hu, B.; Fang, Y.; and Shi, C. 2019. Adversarial Learning on Heterogeneous Information Networks. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019*, 120–129.

Hu, Z.; Dong, Y.; Wang, K.; and Sun, Y. 2020. Heterogeneous Graph Transformer. In *WWW '20: The Web Conference 2020*, 2704–2710.

Jia, D.; Romić, I.; Shi, L.; Su, Q.; Liu, C.; Liu, J.; Holme, P.; Li, X.; and Wang, Z. 2025. Social networking agency and prosociality are inextricably linked in economic games. *Nature Human Behaviour*.

Jin, W.; Ma, Y.; Liu, X.; Tang, X.; Wang, S.; and Tang, J. 2020. Graph Structure Learning for Robust Graph Neural Networks. In *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 66–74.

Jing, B.; Park, C.; and Tong, H. 2021. HDMI: High-order Deep Multiplex Infomax. In *WWW '21: The Web Conference 2021*, 2414–2424.

Kipf, T. N.; and Welling, M. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *5th International Conference on Learning Representations, ICLR 2017*.

Lao, D.; Yang, X.; Wu, Q.; and Yan, J. 2022. Variational Inference for Training Graph Neural Networks in Low-Data Regime through Joint Structure-Label Estimation. In Zhang, A.; and Rangwala, H., eds., *KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 824–834.

Li, K.; Liu, Y.; Ao, X.; Chi, J.; Feng, J.; Yang, H.; and He, Q. 2022. Reliable Representations Make A Stronger Defender: Unsupervised Structure Refinement for Robust GNN. In *KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 925–935.

Liu, N.; Wang, X.; Wu, L.; Chen, Y.; Guo, X.; and Shi, C. 2022a. Compact Graph Structure Learning via Mutual Information Compression. In *WWW '22: The ACM Web Conference 2022*, 1601–1610.

Liu, Y.; Zheng, Y.; Zhang, D.; Chen, H.; Peng, H.; and Pan, S. 2022b. Towards Unsupervised Deep Graph Structure Learning. In *WWW '22: The ACM Web Conference 2022*, 1392–1403.

Lv, Q.; Ding, M.; Liu, Q.; Chen, Y.; Feng, W.; He, S.; Zhou, C.; Jiang, J.; Dong, Y.; and Tang, J. 2021. Are we really making much progress?: Revisiting, benchmarking and refining heterogeneous graph neural networks. In *KDD '21: The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 1150–1160.

Mo, Y.; Nie, F.; Hu, P.; Shen, H. T.; Zhang, Z.; Wang, X.; and Zhu, X. 2024. Self-Supervised Heterogeneous Graph Learning: a Homophily and Heterogeneity View. In *The Twelfth International Conference on Learning Representations, ICLR 2024*.

Park, C.; Kim, D.; Han, J.; and Yu, H. 2020. Unsupervised Attributed Multiplex Network Embedding. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference*, 5371–5378. AAAI Press.

Teru, K. K.; Denis, E. G.; and Hamilton, W. L. 2020. Inductive Relation Prediction by Subgraph Reasoning. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020*, 9448–9457.

Tian, Y.; Dong, K.; Zhang, C.; Zhang, C.; and Chawla, N. V. 2023. Heterogeneous Graph Masked Autoencoders. In *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023*, 9997–10005.

Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; and Bengio, Y. 2018. Graph Attention Networks. In *International Conference on Learning Representations*.

Wang, R.; Mou, S.; Wang, X.; Xiao, W.; Ju, Q.; Shi, C.; and Xie, X. 2021a. Graph Structure Estimation Neural Networks. In *WWW '21: The Web Conference 2021, Virtual Event / Ljubljana, Slovenia, April 19-23, 2021*, 342–353.

Wang, X.; Dong, Y.; Jin, D.; Li, Y.; Wang, L.; and Dang, J. 2023a. Augmenting Affective Dependency Graph via Iterative Incongruity Graph Learning for Sarcasm Detection. In *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence*, 4702–4710.

Wang, X.; Ji, H.; Shi, C.; Wang, B.; Ye, Y.; Cui, P.; and Yu, P. S. 2019. Heterogeneous Graph Attention Network. In *The World Wide Web Conference, WWW 2019, San Francisco, CA, USA, May 13-17, 2019*, 2022–2032.

Wang, X.; Liu, N.; Han, H.; and Shi, C. 2021b. Self-supervised Heterogeneous Graph Neural Network with Co-contrastive Learning. In *KDD '21: The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 1726–1736.

Wang, X.; Wang, Y.; He, D.; Yu, Z.; Li, Y.; Wang, L.; Dang, J.; and Jin, D. 2025. Elevating Knowledge-Enhanced Entity and Relationship Understanding for Sarcasm Detection. *IEEE Trans. Knowl. Data Eng.*, 37(6): 3356–3371.

Wang, Z.; Li, Q.; Yu, D.; Han, X.; Gao, X.; and Shen, S. 2023b. Heterogeneous Graph Contrastive Multi-view Learning. In *Proceedings of the 2023 SIAM International Conference on Data Mining, SDM 2023*, 136–144.

Wu, Q.; Zhao, W.; Li, Z.; Wipf, D. P.; and Yan, J. 2022. NodeFormer: A Scalable Graph Structure Learning Transformer for Node Classification. In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022*.

Yan, F.; Wang, X.; He, D.; Wang, L.; Dang, J.; and Jin, D. 2025. HeterGP: Bridging Heterogeneity in Graph Neural Networks with Multi-View Prompting. In *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence*, 21895–21903.

Yu, D.; Zhang, R.; Jiang, Z.; Wu, Y.; and Yang, Y. 2020. Graph-Revised Convolutional Network. In *Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2020*, 378–393.

Yun, S.; Jeong, M.; Kim, R.; Kang, J.; and Kim, H. J. 2019. Graph Transformer Networks. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019*, 11960–11970.

Zhao, J.; Wang, X.; Shi, C.; Hu, B.; Song, G.; and Ye, Y. 2021. Heterogeneous Graph Structure Learning for Graph Neural Networks. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021*, 4697–4705.

Zhao, J.; Wen, Q.; Ju, M.; Zhang, C.; and Ye, Y. 2023. Self-Supervised Graph Structure Refinement for Graph Neural Networks. In Chua, T.; Lauw, H. W.; Si, L.; Terzi, E.; and Tsaparas, P., eds., *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, WSDM 2023*, 159–167.

Zhou, Z.; Zhou, S.; Mao, B.; Zhou, X.; Chen, J.; Tan, Q.; Zha, D.; Feng, Y.; Chen, C.; and Wang, C. 2023. OpenGSL: A Comprehensive Benchmark for Graph Structure Learning. In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023*.